Structural and functional analysis of genes with potential involvement in resistance to coffee leaf rust: a functional marker based approach

Geleta Dugassa Barka[1,4], Eveline Teixeira Caixeta[1,2,*], Sávio Siqueira Ferreira,[1], Laércio Zambolim[3]

[1]Laboratório de Biotecnologia do Cafeeiro (BIOCAFÉ), BIOAGRO, Universidade Federal de Viçosa (UFV), 36570-000, Viçosa, MG, Brazil

[2]Embrapa Café, Empresa Brasileira de Pesquisa Agropecuária, 70770-901, Brasília, DF, Brazil

[3]Departamento de Fitopatologia, Universidade Federal de Viçosa (UFV), 36570-000, Viçosa, MG, Brazil

[4]Applied Biology Department, Adama Science and Technology University (ASTU), 1888, Adama, Oromia, Ethiopia

[*]Corresponding author

E-mail: eveline.caixeta@embrapa.br (ETC)

Abstract

Physiology-based differentiation of S_H genes and *Hemileia vastatrix* races is the principal method employed for the characterization of coffee leaf rust resistance. Based on the gene-for-gene theory, nine major rust resistance genes (S_H1-9) have been proposed. However, these genes have not been characterized at the molecular level. Consequently, the lack of molecular data regarding rust resistance genes or candidates is a major bottleneck in coffee breeding. To address this issue, we screened a BAC library with resistance gene analogs (RGAs), identified RGAs, characterized and explored for any S_H related candidate genes. Herein, we report the identification and characterization of a gene (gene 11), which shares conserved sequences with other S_H genes and displays a characteristic polymorphic allele conferring different resistance phenotypes. Furthermore, comparative analysis of the two RGAs belonging to CC-NBS-LRR revealed more intense diversifying selection in tomato and grape genomes than in coffee. For the first time, the present study has unveiled novel insights into the molecular nature of the S_H genes, thereby opening new avenues for coffee rust resistance molecular breeding. The characterized candidate RGA is of particular importance for further biological function analysis in coffee.

Introduction

Coffee is one of the most valuable cash crops in many developing economies as it provides employment opportunities in cultivation, processing and marketing activities, thereby sustaining the livelihoods of millions around the world [1]. *H. vastatrix*, the causative agent of coffee leaf rust, accounts for one of the major threats to coffee production in almost every coffee producing region. Despite the release of some resistant coffee cultivars in recent years, coffee rust continues to adversely affect coffee production and undermines the incomes of many households [2]. To date, at least 49 characterized physiological races of *H. vastatrix* have been reported [2,3]. The consistent emergence of new races and the sporadic outbreaks of this disease have imposed challenges in resistance breeding. The most pressing concern is, however, the breakdown of resistance genes leading to disease susceptibility of cultivars that were once validated as superior genetic material for resistance breeding [4].

The molecular profiles of coffee genes involved in different metabolic pathways, their evolution and annotation have been unveiled with the complete sequencing of *C. canephora* genome [5]. Given that *C. canephora* contributes to half of the Arabica coffee genome, being a natural hybrid of *C. canephora* and *C. eugenioides*, open access to its genome has provided valuable insights into the genome of *C. arabica* during the past five years. The discovery and successful introgression of S_H3 resistance gene locus to cultivated Arabica coffee from *C. liberica* was another landmark often considered as one of the greatest milestones in the development of coffee rust resistance [6]. Since then, molecular and physical mapping has enabled the sequencing and annotation of S_H3 region, resulting in the discovery of multiple resistance (R) genes [6,7]. Dominantly inherited, the largest class of R-genes encode nucleotide-binding site leucine-rich-repeat (NBS-LRR) proteins that directly recognize the corresponding virulence (v) protein of the pathogen or its effects [8,9]. These genes are believed to contain several hundred gene families, which are unevenly distributed in the genomes of different plant species [10]. Intracellular signaling domain, similar to Drosophila toll/mammalian interleukin-1 receptor (TNL, Toll-NBS-LRR) and the coiled-coil (CNL, CC-NBS-LRR), are the two major N-terminal amino acid sequences preceding the NBS domain involved in specific signal transduction [8,11]. The other N-terminal domain linked to LRR includes leucine-zipper (a transmembrane protein, TM), protein kinase (PK) and WRKY TIR proteins [12]. These domains are predominantly involved in resistance signal transduction via conformational changes [13]. On the carboxyl-terminal region is the LRR, mediating specific protein-protein interaction to recognize pathogen effectors [14,15]. Although these domains are few in number, nucleotide polymorphism and variability of the LRR region are responsible for the perception of a specific pathogen effector [9,16]. Inter and intraspecific extreme variabilities of NBS-LRR have been attributed to gene duplication, unequal crossing over, recombination, deletion, point mutation and selection pressure due to continuous response to diverse pathogen races [6].

The readily available Arabica coffee BAC libraries constructed from disease resistant genotypes at different laboratories have accelerated studies involving resistance gene cloning [17,18]. Furthermore, the application of arbitrary DNA-based and functional (gene) markers in gene cloning has benefitted crop improvement, either through map-based cloning using the former or direct gene cloning using the latter or both [19]. Direct cloning of the gene of interest over map-based gene cloning is appealing as this method is more precise and straightforward for gene characterization.

In coffee, the origin and organization of disease resistance genes have begun to emerge in recent years as part of an effort to understand the role of major rust resistance genes. One such endeavor was the assembly of R genes spanning the S_H3 locus with the objective of tracing the evolution and diversity of LRR domains in three coffee species [6]. Despite the partial sequencing and annotation of several disease resistance genes in Arabica coffee [20], completely sequenced and characterized candidate genes are not yet readily available. Resistance to rust is conferred by nine major genes (S_H1-9) and the corresponding $v_{1\text{-}9}$ pathogen factors are known for long in the coffee rust pathosystem [3,21]. Nonetheless, molecular and functional characterization of any of the S_H genes and the associated regulatory elements is entirely obscure, yet holds immense potential in changing the perspective of rust resistance breeding. Likewise, the use of functional markers that serve as a direct rust resistance screening tool amongst the host-differential coffee clones is important but is barely addressed. The lack of a typical candidate rust resistance gene is one of the bottlenecks in coffee breeding. A resistance gene analog (RGA) marker, CARF005, was previously confirmed to share disease resistance ORF region in coffee [20,22]. This polymorphic RGA marker encodes the disease resistance protein domain NB-ARC (nucleotide binding site-ARC: ARC for APAF-1, R protein and CED-4 [23], exclusive in coffee cultivars resistant to *H. vastatrix* [22]. The complete sequencing and molecular characterization would help identify candidate disease resistance genes. The state-of-the-art bioinformatics analysis, availability of differential coffee clones with specific S_H genes, structural and functional analysis of conserved domains and associated motifs of candidate RGAs belonging to the S_H gene series could greatly advance coffee rust resistance breeding. Therefore, the objectives of this study were to trace the origin of resistance gene analogs (RGAs) involved in coffee rust resistance and perform comparative molecular characterization of selected candidate gene to determine whether it belongs to the S_H gene series. We also investigated if any of the RGAs were activated during incompatible interaction between C. *Arabica* and *H. vastatrix*..

Materials and methods

Plant materials

Twenty-one differential coffee clones containing at least one of the coffee rust resistance genes (S_H1-9) and three genotypes susceptible to all the virulence factors (v1-9) of *H. vastatrix* were used in the CARF005 screening. The differential clones were initially characterized by CIFC (Centro de Investigação das Ferrugens do Cafeeiro, Portugal) for the identification of the different physiological races of *H. vastatrix*. All clones were vegetatively propagated at the Plant Pathology Department greenhouse of the Universidade Federal de Viçosa (Brazil). Genomic DNA was extracted from a young, second pair of leaves following [24]. DNA integrity was checked by electrophoresing in 1% gel and visualized after staining with ethidium bromide (0.5 µg/ml) and Nanodrop (NanoDrop Technologies, Wilmington, USA). DNA was stored at -20 °C until further use. RNA-Seq libraries (hereafter, referred to as transcriptome) were constructed at 12 and 24 h after infection (hai) during the *C. arabica* CIFC 832/2 –*H. vastatrix* (race XXXIII) incompatible interaction [25] and were used as the reference in the search for novel candidate resistance genes.

PCR conditions

A Sigma made (Sigma-Aldrich, Belo Horizonte, Brazil) disease RGA primer pair, CARF005, (F: 5’-GGACATCAACACCAACCTC-3’ and R: 5’-ATCCCTACCATCCACTTCAAC-3’) [26] was used to screen the differential host clones. PCR reagents were 1x buffer, 0.2 mM dNTPs, 0.2 µM primers, 1 mM $MgCl_2$, 0.8 units of Taq polymerase (Invitrogen, Carlsbad, USA) to which 5 ng gDNA was added to form a reaction volume of 20 µl. PCR cycling parameters were as follows: DNA denaturation at 95°C for 5 min followed by 35 cycles of 94°C for 30 s, 60°C for 30 s and 72°C for 1 min, followed by an extension step at 72°C for 10 min. PCR products were screened for target inserts by electrophoresing in 1% UltraPure^{TM} agarose (Invitrogen) and visualized after staining with ethidium bromide (0.5 µg/ml). All PCR and gel electrophoresis conditions were maintained consistently throughout the study unless stated otherwise.

Screening of BAC clone

BAC library comprising 56,832 clones, constructed using renowned rust resistant Hibrido de Timor clone CIFC 832/2 [18] was used as the target source for RGA (CARF005). These clones were replicated in 384-well titer plates using a plate replicator sterilized with 10% H_2O_2 for 2 min, rinsed in sterile water for 10 seconds and soaked in 70% ethanol in laminar airflow cabinet. After the alcohol evaporated (3-5 min), the old cultures were copied to a new 384-well titer plate containing 70 µl fresh LB media (supplemented with 12.5 µgml^{-1} chloramphenicol) in each well. Culture multiplication was achieved by incubating the plates at 37 °C for 18 h and shaking at 180 rpm. The identification of clones using the CARF005 insert was performed by grouping and subsequent group decomposition of the 384 clones until a single clone was identified as outlined in S1 Fig. BAC DNA was extracted using the centrifugation protocol of Wizard® SV Plus Minipreps DNA Purification System (Promega, Fitchburg, USA).

Sequencing and contig assembly

The single BAC clone isolated using the CARF005 fragment was sequenced using the Illumina HiSeq2000/2500 100PE (paired-end reads) platform at Macrogen (Seoul, South Korea). Paired-end sequence processing and contig assembly were done using SPAdes software [27]. Contigs that matched bacterial genome *(E. coli)* and sequences of the flanking vector (pCC1BAC™) were excluded prior to any downstream sequence processing. The assembled BAC contigs were used to map against a transcriptome constructed from coffee genes that were activated in response to *H. vastatrix* infection by Tophat 2 [28] and to locate the contig region with active gene expression.

Gene prediction and annotation

Contigs with ≥ 200 bp size and sharing ≥ 90% identity with *C. canephora* were subjected to Augustus gene prediction [29]. Among the available genomes in the Augustus dataset, *Solanum lycopersicum* was used as a reference genome, as they shared common gene repertoires and had similar genome sizes [30]. The predicted ORFs were annotated using different online annotation tools. First, BLASTp (NCBI) was used to detect the conserved domains and retrieve their description, followed by the use of Predict Protein Server tool [31] molecular analysis and associated GO search. Protein 3D structure and nucleotide (ATP/ADP/GTP/GDP) binding sites were predicted using I-TASSER suite online tool [32]. As an annotation complement, the predicted ORFs were queried against the database to check for coding sequences (CDS) of *S. lycopersicum* (Sol Genomics Network: https://solgenomics.net/tools/blast/) and *V. vinifera* (Phytozome 11: https://phytozome.jgi.doe.gov) genomes.

Sequence alignment and comparative analysis

Genes encoding resistance proteins were mapped to the *C. canephora* genome [5] to trace their probable origin and organization. BLASTn program was used to query the obtained sequences against the *C. canephora* genome (http://coffee-genome.org/blast). The transcriptome reads (differentially expressed against *H. vastatrix*) were aligned to contig 9 as per Tophat2 (-N 3 --read-gap-length 3 --read-edit-dist 6 --no-coverage-search --b2-very-sensitive) [33] and to locate the region of the contig containing the genes encoded in response to pathogenicity. The intergenic physical position, distance and orientation were analyzed for the RGAs.

Point mutation analysis

The RGAs were analyzed for indels and substitutions using the EMBL MUSCLE multiple sequence alignment tool (http://www.ebi.ac.uk/Tools/msa/muscle/) and MEGA7 [34]. Gene duplication was exclusively analyzed using MEGA 7, while DNA polymorphism and non-synonymous/synonymous substitution rates (ka/ks) were analyzed using DnaSP v5.1 [35].

Functional and phylogenetic analysis

Based on the molecular evolution of protein domains, functional diversity between two NBS-LRR RGAs from coffee was analyzed. Homology was compared for the two RGAs and to identify orthologous genes in the genomes of *S. lycopersicum* (https://solgenomics.net/tools/blast/) and *V. vinifera* (https://phytozome.jgi.doe.gov/pz/portal.html). Subsequently, a protein sequence-based comparative phylogenetic tree was constructed for the two genes and their orthologs from the two related genomes using MEGA7 program [34]. The evolutionary history was inferred using the minimum evolutionary method [36].

Results

Resistance gene screening among the differential coffee clones

To investigate the linkage of RGAs to known SH genes, differential coffee clones with different SH genes were subjected to RGA screening using the functional marker, CARF005. Of the 21 differential coffee clones, the marker was detected in eight clones as presented in Table 1 and S2 Fig. All clones with the S_H6 gene had the marker, while those without the gene failed to amplify in the PCR. Thus, gel analysis of the PCR amplicon revealed that this particular RGA marker amplified the S_H6 gene locus; however, two exceptions were observed. One of them was that the gene's allele was detected in CIFC 128/2-Dilla & Alghe, which is supposed to have just the S_H1 gene. In addition, CARF005 was found to be amplified in a differential-host clone CIFC 644/18 H. Kawisari, for which no S_H gene has been reported to date.

Table 1. S_H gene allelic polymorphism detection in 22 differential coffee clones using CARF005 marker.

No.	Differential clone*	Susceptible to (*H. vastatrix* physiological race)	S_H gene conferred**	Allelic difference (+/-)
1	832/1- Híbrido Timor	-	6,7,8,9,?	+
2	HW17/12	XVI,XXIII	1,2,4,5	-
3	1343/269- Híbrido Timor	XXII,XXV,XXVI,XXVII,XXVIII,XXIX, XXXI,XXXII,XXXIII,XXXVII,XXXIX,XL	6	+
4	H153/2	XII, XVI	1,3,5	-
5	H420/10	XXIX	5,6,7,9	+
6	110/5-S 4 Agaro	X,XIV,XV, XVI,XXIII,XXIV,XXVI, XXVIII	4,5	-
7	128/2-Dilla and Alghe	III, X, XII, XVI, XVII, XIX,XXIII, XXVII	1	+
8	134/4-S12 Kaffa	X, XVI, XIX, XX, XXIII ,XXVII,	1,4	-
9	H419/20	XXIX, XXXI	5,6,9	+
10	635/3-S 12 Kaffa	X, XIV,XV,XVI,XIX, XXIII,XXIV,XXVI,XXVII,XXVIII	1,4,5	-
11	87/1-Geisha	III, X, XII, XVI, XVII, XXIII	1,5	-
12	1006/10-KP 532	XII,XVI,XVII, XXIII	1,2,5	-
13	7963/117-Catimor	XXXIII	5,7 or 5,7,9	-
14	H420/2	XXIX, XXX	5,8	-
15	4106	-	5,6,7,8,9,?	+
16	644/18 H. Kawisari	XIII	?	+
17	832/2- Híbrido Timor	-	6,7,8,9,?	+
18	H147/1	XIV, XVI	2,3,4,5	-
19	32/1-DK1/6	I,VIII, XII, XIV, XVI, XVII, XXIII,XXIV, XXV, XXVIII, XXXI	2,5	-

20	H152/3	XIV,XVI, XXIII, XXIV, XXVII	2,4,5	-
21	33/1-S.288-23	VII, VIII, XII, XIV,XVI,	3,5	-
22	Caturra (c)	All	5	-
23	Catuaí 2143-236 (c)	All	5	-
24	Mundo Novo -376/4 (c)	All	5	-

*Differential clones were from CIFC (Centro de Investigação das Ferrugens do Cafeeiro, Portugal).

**S_H1-9 genes as inferred by CIFC.

Unknown race (-), coffee genotypes used as negative control (c), presence/absence of allelic differences among S_H genes (+/-) and unknown S_H gene (s) (?).

Sequence analysis of ORFs

Identification of a BAC clone using CARF005 and the comparative analysis of the RGAs with the other ORFs from *C. canephora* was performed to localize their relative position and determine the putative function. To characterize genetic loci conferring resistance to leaf rust, a BAC clone 78-K-10 (with ~146 kb insert) was identified as shown in S1 b & c Fig. Illumina HiSeq2000/2500 100PE generated 8,711,320 reads. After removing vector sequences and noisy sequences, quality sequences (>20 QC) were assembled into 86 contigs of ≥ 200bp from which the two contigs, contigs 3 (16570 bp) and 9 (8285 bp), were selected (as they had >90% similarity with *Coffea canephora* DNA sequence) and then subjected to downstream processing. The sequences of the two contigs were combined and deposited at NCBI (accession number: KY485942). These contigs shared ≥ 90% identity with *C. canephora* contigs at different chromosomal regions with the highest identity (99% for contig 3 and 97% for contig 9) being on chromosome 0. All the 13 ORFs predicted (eight in contig 3 and five in contig 9) had matched to different species when queried against BLASTp database or to the *C. canephora* genome hub with significant similarities (≤ 1e^{-05} e-value) when BLASTn was used as presented in S3 Table. Among these, five genes (genes 5, 9, 10, 11 and 12) shared significant identities with RGAs from *C. canephora*. These genes are homologous to sequences in the *C. canephora* genome with the highest query coverage being on chromosome 0 as presented in Table 2. Genes 5 (intron-less, 1130 aa) and 11 (with two introns and two exons, 1118 aa) were the largest genes predicted. Both genes were located on the negative reading frames that belong to the CC-NBS-LRR gene family. Mapping to the *C. canephora* genome showed that these genes are separated by 1,634,522 bp, although they are delimited with a shorter length (460 bp) in *C. arabica*. In contrast, the other four RGAs matched and retained their expected positions in the *C. arabica* genome as shown in S4 Fig.

Table 2. Size and structure of five resistance gene analogs and their mapping to chromosome 0 of *C. canephora* genome.

	Genes*				
	5	9	10	11	12
Contig	3	9	9	9	9
Exon 1	3393	113	121	1175	345
Intron 1	-	554	87	611	1786
Exon 2	-	118	112	2222	183
Intron 2	-	121	711	124	-
Exon 3	-	69	121	-	-
Intron 3	-	91	-	-	-
Exon 4	-	155	-	-	-
Intron 4	-	476	-	-	-
Exon 5	-	130	-	-	-
Query coverage (%)	99.94	72.68	30.48	99.46	97.33
Identity (%)	76.00	85.00	79.00	68.84	73.00
E-value	0.00	9,00E-30	5,00E-17	0.00	3,00E-48
Frame	N	N	P	N	P
Start hit-End hit	108638370-108641761	106998076-106999730	107000654-107000761	107000357-107003848	107000234-107004551
Protein (aa)	1130	194	117	1118	175

*Exon and intron sizes are in nucleotides.

N, negative reading frame and P, positive reading frame.

Gene prediction was performed by Augustus command-line version gene prediction [29].

CARF005 amplicon verification

Web-based PCR analysis was conducted to validate the specificity of the CARF005 primer pair and to complement the PCR amplification experiments. *In silico* PCR analysis using contig 9 and gene 11 ORF as the template strands, indicated that the CARF005 marker had a size of 400 bp (http://www.bioinformatics.org/sms2/pcr_products.html). This size of the amplicon was confirmed by PCR using the template DNA from the clone 78-K-10 as outlined in S1 c Fig. Notably, the amplicon spans from position 6867 to 7266 bp in contig 9 (8285 bp) in a negative orientation and from position 2065 to 3115 bp in the ORF of gene 11(3354 bp), respectively.

Gene annotation

Gene annotation was performed to identify the putative protein-coding gene. The annotation of 13 ORFs showed a range of protein arrays most of which had no role in disease resistance and lacked conserved domains. Among the five RGAs detected in either NCBI BLASTp, or BLASTn against the *C. canephora* genome, genes 9 (unnamed protein product), 10 (putative resistance gene) and 12 (putative resistance gene) had similarity to RGAs as observed by mapping to *C. canephora* genome as presented in S3 Table. Yet, genes 5 and 11 encoded the largest resistance proteins (Gene 5: 126.81 kDa and pi: 7.65; gene 11: 126.67 kDa and pi: 8.44) and identified several resistance associated GO terms characterizing their multiple functional domains as shown in Table 3.

Table 3. Annotation and functional comparison of gene 5 and 11.

Molecular function ontology				
GO ID	**GO term**	**Reliability (%)**	**Gene 5**	**Gene 11**
GO:1901363	Heterocyclic compound binding	49	✓	✓
GO:0000166	Nucleotide binding	49	✓	✓
GO:0005488	Binding	49	✓	✓
GO:1901265	Nucleoside phosphate binding	49	✓	✓

GO:0097159	Organic cyclic compound binding	49	✓	✓
GO:0036094	Small molecule binding	49	✓	✓
GO:0097367	Carbohydrate derivative binding	41	✓	✓
GO:0017076	Purine nucleotide binding	41	✓	✓
GO:0032559	Adenyl ribonucleotide binding	41	✓	✓
GO:0032555	Purine ribonucleotide binding	41	✓	✓
Biological process ontology				
GO:0006952	Defense response	36	✓	✓
GO:0006950	Response to stress	36	✓	✓
GO:0050896	Response to stimulus	36	✓	✓
GO:0002376	Immune system process	16	✓	✓
GO:0006955	Immune response	16	✓	✓
GO:0045087	Innate immune response	16	✓	✓
GO:0044699	Single-organism process	14	✓	✓
GO:0009987	Cellular process	14	✓	✓
GO:0044763	Single-organism cellular process	14	✓	✓
GO:0033554	Cellular response to stress	12	✓	✓
GO:0016265	Death	12	✓	✓
GO:0051716	Cellular response to stimulus	12	✓	✓
GO:0012501	Programmed cell death	12	✓	✓
GO:0008219	Cell death	12	✓	✓
GO:0034050	Host programmed cell death induced by symbiont	12	✓	✓
GO:0009626	Plant-type hypersensitive response	12	✓	✓
GO:0009814	Defense response, incompatible interaction	7	✓	✓
Cellular component ontology				
GO:0016020	Membrane	33	✓	✓
GO:0044464	Cell part	33	✓	✓
GO:0005623	Cell	33	✓	✓

GO:0005737	Cytoplasm	32	✓	✓
GO:0044424	Intracellular part	32	✓	✓
GO:0005886	Plasma membrane	31	✓	✓
GO:0071944	Cell periphery	31	✓	✓
GO:0043227	Membrane-bounded organelle	24	✓	✓
GO:0043226	Organelle	24	✓	✓
GO:0005634	Nucleus	24	✓	✓

Annotation was performed by Predict Protein online server [31] (URL: https://www.predictprotein.org).

Gene characterization

To identify the candidate R genes activated against *H. vastatrix* incursion, two contigs (contigs 3 and 9) were mapped against the transcriptome of *C. arabica*-*H. vastatrix* interaction [25]. Genes 9, 10, 11 and 12 were mapped to transcripts differentially-expressed during incompatible interactions at 12 and 24 hai with *H. vastatrix* (race XXXIII) as illustrated in S5 Fig. Contig 3, from which gene 5 was predicted, was also mapped against the same transcriptome resulting in no matching transcripts that were differentially expressed at the two time points (12 and 24 hai) following pathogen inoculation. Contig 9 region, where most R genes are positioned was estimated by mapping against the transcriptome. The result showed approximately 81.58% of the contig encodes transcripts of varying lengths associated with rust resistance, which are activated at 12 and 24 hai in response to *H. vastatrix* inoculation. Further analysis of genes 5 and 11 revealed that they belong to the NBS-LRR gene family (the major R genes in plants), suggesting the importance of continuing the *in silico* comparative structural and functional analysis. Intriguingly, both have the Rx-cc-like coiled-coil potato virus x resistance protein domain and four additional multi-domains featuring the entire protein sequence (Fig 1). These genes can be referred as CC-NBS-LRR, as they comprise the N-terminal CC and LRR C-terminal domains flanking the NBS on either side.

Fig 1. Comparison of conserved domains and motif architecture in genes 5 and 11. Note the polymorphism of domains in both genes (encircled by red spheres). Conserved domains were detected using NCBI BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) database.

In addition, annotation of both genes indicates that they encode defense proteins involved in various biological defense as demonstrated in Table 3. Notably, although these genes share 90.24% nucleotide identity, their amino acid sequence identity is only 80.03% (Fig 2). The possibility of substitution mutation events was considered in explaining the diversity. Accordingly, the overall amino acid diversity was attributed to non-synonymous substitution events (non-synonymous/synonymous ratio, ka/ks = 1.5913) in both genes. Further analysis of LRR region showed a higher rate of non-synonymous substitution mutation (ka/ks, non-synonymous/synonymous substitution ratio = 1.9660).

Fig 2. Alignment of proteins encoded by genes 5 and 11 and the protein binding regions. *In silico* prediction of protein binding regions of gene 5 (boxed in red), protein binding regions of gene 11 (boxed in green) and conserved protein binding region (boxed in blue) are shown. Amino acid substitution: unrelated amino acid substitution (space), weakly similar substitution (period), strongly similar substitution (colon) and conserved amino acids (star, unmarked). Note the seven substitution mutations resulting in polymorphism of seven protein binding sites (purple encircled) in either of the sequence. Sequence alignment was carried out using Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/).

Comparative analysis of structural and functional sites

Structural modeling and comparative analyses of the identified genes was performed in order to infer the possible protein functions. Therein, we found that the number of LRR domains in genes 5 and 11 is conserved (13 repeats in both), but arranged differently. We noted the introduction of a coenzyme domain (CoaE, dephospho-CoA kinase) in gene 5, while a LRR variant (LRR_8) seemed to be introduced in gene 11 (Fig 1). Despite sharing similar protein multi-domains, two trans-membrane motifs (spanning 5-22 and 102-119 amino acid regions) were detected exclusively in the coiled-coil domain of gene 11 (data not shown). The amino acid sequences of genes 5 and 11 were further analyzed for protein and nucleotide binding site polymorphism. Protein binding sites of the two genes revealed different sensitivity to substitution mutations. Few sites that were specific to each gene were highly affected while most of the binding sites had moderate to no effect as presented in S6 Table. The analysis revealed 17 protein binding sites in gene 5 and 14 sites in gene 11. Similarly, their secondary structures and solvent accessibility properties revealed conserved features (Fig 3aI-IV & cI-IV). Nevertheless, the amino acid residues forming the protein binding sites of the two genes showed high variability in the LRR regions (Fig 2; Fig 3a & c). Although most of the residues are not conserved, the ADP binding site of the NBS domain contained some conserved sites (Fig 3b.II & d.II).

Fig 3. *In silico* 3D structure, protein and nucleotide binding site prediction for gene 5 (A and B) and 11 (C and D). Protein binding and secondary structure (A and C): Protein binding sites (I), the three types of secondary structures are assumed at different regions (helical: red boxes, strand: blue boxes and loop: intervening white spaces) (II), solvent accessibility (exposed: blue boxes, buried: yellow boxes and intermediate: white spaces) (III) and high protein disorder and flexibility: green boxes (IV). 3D structure and nucleotide binding sites (B and D): 3D structures with Rx-CC-like (blue) to LRR (red to light: orange forming the 'horseshoe' structure) domains are as shown in Figs 3b.I & d.I and the colored residues (NBS) forming the nucleotide (ATP/GTP/ADP/GDP)-binding site (BII and DII). Nucleotide binding site residues with the highest C-score are listed in the right box (conserved residues highlighted in yellow) with the red arrow indicating the sites. I-TASSER modelling C-score [32] was -1.73 and -1.75 (C-score ranging from -5 to 2, where 2 refers to the highest confidence) and 0.29 and 0.17 (C-score ranging from 0-1, where higher score indicates reliable prediction) for nucleotide binding prediction for the two proteins, respectively.

Interlocus comparison of S_H genes

To investigate the conserved regions in the five RGAs (genes 5, 9, 10, 11 and 12), contigs 3 and 9 were queried against three *C. canephora* and 10 *C. arabica* specific contigs assembled from BAC clones spanning S_H3 locus from the work of [6]. All the 10 S_H3 contigs matched with contig 3 but with varying alignment lengths and identities as presented in S7 Table. Contig GU123898 and HQ696508 (both specific to *C. arabica*) had the highest number of matches to contig 9 (from which four clustered RGAs were predicted) with alignment lengths of 170 nts (77.647% identity and $1.57e^{-31}$ e-value) and 179 nts (76.536% identity and

1.21e^{-26} e-value), respectively. The closest contig (HQ696508) is located on the complementary strand of gene 11 and is 505 bp upstream of the position where CARF005 forward primer annealed to gene 11.

Phylogenetic analysis

In an attempt to discern the ancestral relationship of a set of sequences, phylogenetic analysis was performed. Accordingly, two resistance gene families (the NBS-LRR and non-NBS-LRR) were identified, completely sequenced and mapped to chromosome 0 of *C. canephora* genome with a query coverage of 99.94% for genes 5 and 11, 72.68% for gene 9, 33.05% for gene 10 and 97.52% for gene 12. The diversity of the NBS-LRR family was detected by analyzing the ka/ks ratios as presented in Table 4. The analysis revealed that the non-synonymous substitution event is common in the CDS, as revealed from all the pairwise analyses. Furthermore, the non-synonymous substitution of CDS is more prominent in the LRR region (in almost all pairwise comparisons) as demonstrated in Table 4.

Table 4. Pair-wise synonymous and non-synonymous nucleotide substitution analysis among the six resistance gene analogs (gene 5 and 11 and their respective two top hits as mined by BLASTn in NCBI).

		Entire protein			LRR region		
Seq. 1	Seq. 2	Ks	Ka	ka/ks	Ks	Ka	ka/ks
gene5_hit1	gene11_hit1	0.0786	0.1302	1.6565	0.0702	0.1383	1.9701
gene5_hit1	gene11_hit2	0.0899	0.1614	1.7953	0.0536	0.1408	2.6269
gene5_hit1	gene11	0.0723	0.1177	1.6279	0.0622	0.1233	1.9823
gene5_hit1	gene5_hit2	0.0635	0.0999	1.5732	0.0583	0.1029	1.7650
gene5_hit1	gene5	0.0009	0.0039	4.3333	0.0015	0.0045	3.0000
gene11_hit1	gene11_hit2	0.1092	0.1854	1.6978	0.0756	0.1602	2.1190
gene11_hit1	gene11	0.0061	0.0164	2.6885	0.0095	0.0170	1.7895
gene11_hit1	gene5_hit2	0.0761	0.1309	1.7201	0.0736	0.1369	1.8601
gene11_hit1	gene5	0.0786	0.1288	1.6387	0.0701	0.1383	1.9729
gene11_hit2	gene11	0.1074	0.1742	1.6220	0.0686	0.1445	2.1064
gene11_hit2	gene5_hit2	0.0846	0.5829	6.8901	0.0607	0.1440	2.3723
gene11_hit2	gene5	0.0902	0.1620	1.7960	0.0540	0.1430	2.6481
gene11	gene5_hit2	0.0704	0.1155	1.6406	0.0616	0.1199	1.9464

gene11	gene5	0.0723	0.1164	1.6100	0.0622	0.1234	1.9839
gene5_hit2	gene5	0.0635	0.0997	1.5701	0.0583	0.1052	1.8045

Seq., sequence, non-synonymous/synonymous substitution rate was computed by DnaSP v5.1 [35].

Moreover, phylogenetic analysis showed that the tomato gene 5 was closely related to genes 5 and 11 of coffee than the gene 11 of both tomato and grape (Fig 4). Within coffee itself, a significant diversity between genes 5 and 11 was detected by the MEGA 7 bootstrap method of the phylogenetic test.

Fig 4. Phylogenetic history of genes 5 and 11 in three related genomes. The evolutionary history was inferred using the Minimum Evolution method [36]. The optimal tree with the sum of branch length = 2.98805978 is shown. The percentage of replicate trees with the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches [67]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [68] and are in the units of the number of amino acid substitutions per site. The ME tree was searched using the Close-Neighbor-Interchange (CNI) algorithm [69] at a search level of 1. The Neighbor-joining algorithm [70] was used to generate the initial tree. The analysis involved 6 amino acid sequences. All positions containing gaps and missing data were eliminated. A total of 554 positions were there in the final dataset. Evolutionary analyses were conducted in MEGA7 [34]. Subject IDs are indicated in parenthesis for the corresponding two homologous sequences mined by BLASTx against tomato (Sol Genomics Network) and grape (Phytozome) genome databases.

Discussion

The majority of NBS-LRR encoding genes are known to be clustered but unevenly distributed in plant genomes [10,37–39]. The NBS-ARC domain is known to be involved in directly blocking the biotrophic pathogens by activating the hypersensitive response (HR) [40]. HR starts with programmed cell death of affected and surrounding cells and ends with the activation of systemic acquired resistance (SAR), in which the defense is induced in distal non-infected cells of the host under attack [41,42]. By recognizing the corresponding virulence (vr) factors or their effects, NBS-LRR proteins are sufficient to induce HR [8,9,42,43]. In the present study, a cluster of two different classes of RGAs resistant to coffee rust, the NBS-LRRs linked to non-NBS-LRR genes were reported. The two NBS-LRR genes (genes 5 and 11) are the largest non-TNL genes sequenced in Arabica coffee and most other plants investigated to date [6,44–46].

We identified 13 genes: genes 1, 2, 3, 4, 5, 6, 7 and 8 (gene 5 is a RGA) from contig 3 and genes 9, 10, 11, 12, and 13 (only 13 is not a RGA) from contig 9. We also report, a completely sequenced and characterized novel RGA (gene 11), from Híbrido de Timor CIFC 832/2, probably a major component of the S_H gene. This Híbrido de Timor (HDT) (*C. arabica* x *C. canephora*) is immune to all known virulence factors of *H. vastatrix* physiological races and therefore, is an extremely important source of resistance [47]. Additionally, the mapping against the transcriptome of *C. arabica*-*H. vastatrix* interaction suggests that the three other clustered RGAs (genes 9, 10 and 12) have differential-expression during the incompatible interaction. Mapping study has also revealed the presence of reads exclusively mapped to transcripts of pathogen-infected plants at 12 and 24 hai.

The known rust resistance genes, S_H3 in *C. liberica* [48], S_H6, 7, 8 and 9 in *C. canephora* [49] and S_H1, 2, 4 and 5 genes in *C. arabica* are dominantly-inherited genes [50]. One of the fundamental questions to be clarified is how different are these 9 S_H genes that belong to different coffee species. The comparative analysis of contigs from the S_H3 locus of *C. arabica* and *C. canephora* [6] revealed different levels of conservation of motifs in the two contigs examined: contigs 3 and 9. The results indicate that the RGAs may share large conserved regions, but few highly polymorphic regions encoding specific protein motifs necessary for critical roles. This characteristic conservation of domains was once more confirmed based on comparative analysis of the cloned gene (gene 11) and using differential coffee clones for S_H gene identification. PCR amplification of gene 11 also indicated the existence of allelic difference/polymorphism among the S_H gene loci and considerable sequences of conserved domain (on which CARF005 primer was designed) with S_H6 and possibly with S_H1. As PCR amplification using CARF005 primer (constituting gene 11) was detected in all the differential clones with S_H6 and 832/1-HT and 832/2-HT containing S_H6, 7, 8, 9 and S_H?. In addition, we report a conserved sequence of gene 11 11 in CIFC 128/2-Dilla & Alghe, previously considered to contain only the S_H1 gene, and in CIFC 644/18 H. Kawisari with an uncharacterized S_H-gene. Overall, we propose the following hypothesis for extensive and rigorous biological investigation: the identified gene (gene 11) could be one of the unidentified and a not yet supplanted (at least in Brazil) S_H gene in HDT consisting of a conserved domain (CARF005) shared with the S_H6 and S_H1 genes.

Mapping of the RGAs to *C. canephora*, the result from differential clone screening and annotation altogether confirmed that gene 11 locus is descended from *C. canephora,* hence is a sibling of S_H6-9 [49]. Besides, mapping of RGAs to *C. canephora* was complemented by the differential clone screening for SH genes, which affirms the existence of strong linkage of SH gene locus and the RGAs as their conserved sequences (CARF005) were detected in eight of the differential clones. The disparity of the position of gene 5 in relation to gene 9 (the fact that all the predicted genes are from an insert size of ~146 kbp) could be attributed to the linked LTR-retrotransposon (GROUP_78_RLC4) as demonstrated in S4 Fig and the transposase gene (gene 1). Transposons could have interrupted and separated the two genes apart by 1.6 Mb, since *C. arabica* is known to have diverged from *C. canephora*. Multiple transposable elements linked to NBS-LRR regions were reported in other plants [45,51]. Transposition of genes and gene fragments are some of the mechanisms that generate variability and positional changes among the NBS-LRR genes in different plants [45,51–54]

Rx-CC, PLN and NB-ARC domains are conserved in the NBS-LRR genes across diverse plant species [44,55,56]. The potato virus x resistance (Rx) protein-like N-terminal coiled-coil domain mediates intramolecular interaction with NB-ARC and intermolecular interactions through RanGAP2 (Ran-GTPase-activating protein-2) in potato [43,57]. Rx-CC, RanGAP2 interaction site and NB-ARC were detected in genes 5 and 11, suggesting similarity in their defense role in coffee. However, unlike the Rx-CC domain with four helical structures, five helical structures are conserved in genes 5 and 11, indicating polymorphic differences between the species. The PLN00113 domain in gene 5 and PLN03210 in gene 11, span the LRR region and were initially reported in *A. thaliana* [44]. The distinct position of these domains in genes 5 and 11 indicates high variability in the LRR region in both genes. Functional motif prediction indicated that the PLN03210 (LRR domain) is likely engaged in direct effector interaction while the corresponding PLN00113 of gene 5 is engaged in LRR-reception and downstream kinase-mediated signaling. These observations are in accordance with the functional and structural analysis data of LRR proteins in *A. thaliana* [44,58–61]. Based of their annotations, the two genes (genes 5 and 11) products are intracellular resistance proteins that directly or indirectly recognize pathogen effector proteins and subsequently trigger a response that may be as severe as localized cell death [42].

Different selection pressures shape the evolution of domains in the NBS-LRR encoding genes. The NBS domain was assumed to be under the purifying selection (a negative selection in which variation is minimized by stabilizing selection) than by the diversifying selection, which acts on the LRR domain [9,62]. In contrast, the diversifying selection (positive selection) act on all the domains of genes 5 and 11 (ka/ks >1). This result is contrary to the general assumption that diversifying selection is diluted when the overall non-synonymous substitution is considered [6], indicating an intense diversifying selection action on both genes. Further investigation of four more orthologous genes also resulted in similar findings, indicating that the NBS-LRR genes are highly variable due to substitution mutations. As the LRR domains are involved in direct ligand binding, their variability due to non-synonymous substitution is higher than that seen in other domains. This results in the formation of a super-polymorphic region to cope with the continuously evolving pathogen effectors. Similar findings (from different plants, including coffee) on diversifying selection have been reported [6,9,11,38,45,63,64]. Diversifying selection by non-synonymous substitution was detected in non-NBS-LRR genes (genes 10 and 12) (data not shown), reiterating the importance of substitution mutation in such clustered R genes. Synergistic activation of the two groups (NBS-LRR and non-NBS-LRR) may enhance the resistance durability; and so their expression pattern merits further investigation.

Based on the phylogenetic tree of orthologous genes originated from related genomes, the six genes could be divided into two groups. Gene 5 from tomato is closely related to genes 5 and 11 from coffee, making the first group, whereas genes 5 and 11 from grape happens to be the second highly diversified group. Intraspecies diversity of non-TIR-NBS-LRR due to substitution and genetic recombination exist in grape [65] and tomato [66] while gene duplication and conversion events were observed in coffee [6]. In general, the phylogenetic tree revealed that genes 5 and 11 may have recently diverged in coffee, while the divergence observed in the other species may have been earlier events.

Conclusion

The two groups of RGAs, NBS-LRR and non-NBS-LRR, are clustered in a single locus from which multiple variants of resistance genes are expressed to confer specific resistance functions. The four cloned, sequenced and characterized RGAs span a rust resistance gene locus descended from *C. canephora*. The two CC-NBS-LRR protein encoding genes are under strong diversifying selection impacting all component domains. A more intense diversification of LRR region indicates that the variability in the effector binding site is the cause of divergence in resistance specificity. Although conserved sequences were detected for the S_H6 gene across the various differential coffee clones, it could be inferred that the S_H gene loci have a characteristic polymorphism conferring different resistance phenotypes against coffee leaf rust. This is the first report unveiling new insights into the molecular nature of S_H genes. The CC-NBS-LRR gene thus characterized is the largest and most complete sequence ever reported in Arabica coffee. The work demonstrated a cluster of resistance genes spanning the R gene locus that could serve as functional markers for subsequent functional analysis. These findings could also serve as a benchmark for validation of expression patterns in response to pathogenicity and gene segregation along generations. Such studies can be applied in molecular breeding as it has the potential to replace arbitrary DNA-based marker-assisted breeding at least for two reasons. First, there is no loss due to segregation, which is the case even for finely saturated markers. Second, four of the RGAs (genes 9, 10, 11 and 12) are stacked in a locus, from which different primers can be designed to screen genotypes to verify co-segregation analysis.

Supporting information

S1 Fig. Work flow in BAC clone screening. Clone pooling and subsequent group decomposition to isolate a single clone with CARF005 insert (A), DNA of isolated clone 78-K-10 (B) and CARF005 PCR amplicon (C) as revealed by 1% UltraPure^{TM} agarose gel electrophoresis. M is 100 bp DNA size marker.

S2 Fig. The 21 differential coffee clones screened for CARF005 marker (listed in order as in Table 1). Clones with CARF005 were 1 (832/1-HT), 3 (1343/269-HT), 5 (H420/10), 7 (128/2-Dilla and Alghe), 9 (H419/20), 15 (4106), 16 (644/18 H. Kawisan, a new report) and 17 (832/2-HT). M: DNA weight marker ladder (the lightest band being 100 bp). No gel cropping was performed to any of the lanes displayed in the gel above.

S3 Fig. Mapping of RGAs clustered on chromosome 0 of *C. canephora*. Putative mRNA transcription orientations are shown by black arrows. The relatively larger size of gene 12 is due to the largest size of its intron 1 (Table 2). Green boxes are used to mark query positions relative to subject genes (other gene products, all in blue boxes). Note that gene 10 and 12 are in positive orientation (Table 2) with no matching transcript here, hence probably originated from different parent or attributed to mutation events in *C. arabica*. Mapping was carried out by CDS (coding sequence) BLASTn followed by track assembly on *C. canephora* genome hub server [5].

S4 Fig. Mapping of contig 9 to transcriptome of differentially expressed genes during *C. arabica-H. vastatrix* (race XXXIII) incompatible interaction to show the region of active gene (gene 9, 10, 11 and 12) expression. Note the three expression profiles (three rows) corresponding to control (uninoculated at 0 hour, top row), 12 (middle row) and 24 hai (bottom row) of transcriptome reads mapped against contig 9 of resistant coffee clone (CIFC HDT 832/2). Grey shades indicate matching transcriptome reads while nucleotide substitutions (mismatches) were shown by colored strips (yellow: G, green: A, red: R and blue: C). Large red shades indicate deletions. Contig mapping was performed by Tophat 2[28] setting alignment parameter as '-N 3 --read-gap-length 3 --read-edit-dist 6 --no-coverage-search --b2-very-sensitive' to locate the region of the contig encoding genes against the pathogen and visualized with Integrative Genomics Viewer (IGV) v. 2.3[33].

S1 Table. Top hits for the 13 ORFs as found in NCBI by BLASTp or at *C. canephora* genome by BLASTn. *Homologous sequences for which no ID/Accession number has been assigned are indicated in hyphen. BLASTp was performed by NCBI online server (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins).

S2 Table. Mutation (substitution) effect on protein binding regions of gene 5 and 11 indicated by amino acid sequence in respective genes. *Hyphen indicates range of amino acids constituting the binding site.

Yellow highlighted residues are conserved residues in both genes while purple highlighted residues are specific protein binding sites in respective gene. Substitution mutation effect analysis was performed by The Predict Protein Server[31].

S3 Table. Output of the two contigs BLASTed against S_H3 locus contigs specific to C. arabica and C. canephora*. *Ten contigs specific to *C. arabica* and three contigs specific to *C. canephora*, all assembled from BAC clones with S_H3 locus were taken from the work of [6].

Acknowledgements

We thank Drs Jorge L.B. Pacheco, Abraham Abera, Bayissa Chala and Mohammed Naimuddin for their valuable suggestions and edition of the manuscript. We are also grateful to the Agronomic Institute of Paraná, Londrina-Brazil, for providing the CIFC 832/2 BAC library.

Author contributions

Conceptualization: GDB ETC

Formal analysis: GDB SSF

Investigation: GDB

Methodology: GDB ETC SSF

Project administration: ETC LZ

Resources: ETC LZ

Software: GDB SSF

Supervision: ETC LZ

Validation: GDB ETC SSF LZ

Writing-Original draft: GDB

Writing-Review & editing: GDB ETC SSF LZ

References

1. Mussatto SI, Machado EMS, Martins S, Teixeira JA. Production, Composition, and Application of Coffee and Its Industrial Residues. Food Bioprocess Technol. 2011;4: 661–672. doi:10.1007/s11947-011-0565-z

2. Zambolim L. Current status and management of coffee leaf rust in Brazil. Trop Plant Pathol. 2016;41: 1–8. doi:10.1007/s40858-016-0065-9

3. Gichuru EK, Ithiru JM, Silva MC, Pereira AP, Varzea VMP. Additional physiological races of coffee leaf rust (Hemileia vastatrix) identified in Kenya. Trop Plant Pathol. 2012;37: 424–427. doi:10.1590/S1982-56762012000600008

4. Cristancho MA, Botero-Rozo DO, Giraldo W, Tabima J, RiaÃ±o-PachÃ³n DM, Escobar C, et al. Annotation of a hybrid partial genome of the coffee rust (Hemileia vastatrix) contributes to the gene repertoire catalog of the Pucciniales. Front Plant Sci. 2014;5. doi:10.3389/fpls.2014.00594

5. Denoeud F, Carretero-Paulet L, Dereeper A, Droc G, Guyot R, Pietrella M, et al. The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. Science (80-). 2014;345: 1181–1184. doi:10.1126/science.1255274

6. Ribas AF, Cenci A, Combes M-C, Etienne H, Lashermes P. Organization and molecular evolution of a disease-resistance gene cluster in coffee trees. BMC Genomics. 2011;12: 240. doi:10.1186/1471-2164-12-240

7. Cenci A, Combes M-C, Lashermes P. Comparative sequence analyses indicate that Coffea (Asterids) and Vitis (Rosids) derive from the same paleo-hexaploid ancestral genome. Mol Genet Genomics. 2010;283: 493–501. doi:10.1007/s00438-010-0534-7

8. Jones JDG, Dangl JL. The plant immune system. Nature. 2006;444: 323–329. doi:10.1038/nature05286

9. McHale L, Tan X, Koehl P, Michelmore RW. Plant NBS-LRR proteins: adaptable guards. Genome Biol. 2006;7: 212. doi:10.1186/gb-2006-7-4-212

10. Hulbert SH, Webb CA, Smith SM, Sun Q. Resistance gene complexes: evolution and utilization. Annu Rev Phytopathol. 2001;39: 285–312. doi:10.1146/annurev.phyto.39.1.285

11. DeYoung BJ, Innes RW. Plant NBS-LRR proteins in pathogen sensing and host defense. Nat Immunol. 2006;7: 1243–1249. doi:10.1038/ni1410

12. Liu J, Liu X, Dai L, Wang G. Recent Progress in Elucidating the Structure, Function and Evolution of Disease Resistance Genes in Plants. J Genet Genomics. 2007;34: 765–776. doi:10.1016/S1673-8527(07)60087-3

13. Leipe DD, Koonin E V., Aravind L. STAND, a Class of P-Loop NTPases Including Animal and Plant Regulators of Programmed Cell Death: Multiple, Complex Domain Architectures, Unusual Phyletic Patterns, and Evolution by Horizontal Gene Transfer. J Mol Biol. 2004;343: 1–28. doi:10.1016/j.jmb.2004.08.023

14. Van der Hoorn RAL. Identification of Distinct Specificity Determinants in Resistance Protein Cf-4 Allows Construction of a Cf-9 Mutant That Confers Recognition of Avirulence Protein AVR4. PLANT CELL ONLINE. 2001;13: 273–285. doi:10.1105/tpc.13.2.273

15. Kushalappa AC, Yogendra KN, Karre S. Plant Innate Immune Response: Qualitative and Quantitative Resistance. CRC Crit Rev Plant Sci. 2016;35: 38–55. doi:10.1080/07352689.2016.1148980

16. Ellis J, Dodds P, Pryor T. Structure, function and evolution of plant disease resistance genes. Curr Opin Plant Biol. 2000;3: 278–284. doi:10.1016/S1369-5266(00)00080-7

17. Combes M-C, Lashermes P, Chalhoub B, Noir S, Patheyron S. Construction and characterisation of a BAC library for genome analysis of the allotetraploid coffee species (Coffea arabica L.). TAG Theor Appl Genet. 2004;109: 225–230. doi:10.1007/s00122-004-1604-1

18. Cação SMB, Silva N V., Domingues DS, Vieira LGE, Diniz LEC, Vinecky F, et al. Construction and characterization of a BAC library from the Coffea arabica genotype Timor Hybrid CIFC 832/2. Genetica. 2013;141: 217–226. doi:10.1007/s10709-013-9720-y

19. Poczai P, Varga I, Laos M, Cseh A, Bell N, Valkonen JP, et al. Advances in plant gene-targeted and

functional markers: a review. Plant Methods. 2013;9: 6. doi:10.1186/1746-4811-9-6

20. Noir S, Combes MC, Anthony F, Lashermes P. Origin, diversity and evolution of NBS-type disease-resistance gene homologues in coffee trees (Coffea L.). Mol Genet Genomics. 2001;265: 654–62. Available: http://www.ncbi.nlm.nih.gov/pubmed/11459185

21. Rodrigues CJ, Bettencourt AJ, Rijo L. Races of the Pathogen and Resistance to Coffee Rust. Annu Rev Phytopathol. 1975;13: 49–70. doi:10.1146/annurev.py.13.090175.000405

22. Alvarenga MS, Caixeta TE, Hufnagel B, Thiebaut F, Maciel-Zambolim E, Zambolim L, et al. In silico identification of coffee genome expressed sequences potentially associated with resistance to diseases. Genet Mol Biol. 2010;33: 795–806. Available: http://www.scielo.br/pdf/gmb/v33n4/31.pdf

23. Van der Biezen EA, Jones JD. Plant disease-resistance proteins and the gene-for-gene concept. Trends Biochem Sci. 1998;23: 454–6. Available: http://www.ncbi.nlm.nih.gov/pubmed/9868361

24. Diniz LEC, Sakiyama NS, Lashermes P, Caixeta ET, Oliveira ACB, Zambolim EM, et al. Analysis of AFLP markers associated to the Mex-1 resistance locus in Icatu progenies. Crop Breed Appl Biotechnol. 2005;5: 387–393. doi:10.12702/1984-7033.v05n04a03

25. Florez JC, Mofatto LS, do Livramento Freitas-Lopes R, Ferreira SS, Zambolim EM, Carazzolle MF, et al. High throughput transcriptome analysis of coffee reveals prehaustorial resistance in response to Hemileia vastatrix infection. Plant Mol Biol. 2017; doi:10.1007/s11103-017-0676-7

26. Alvarenga M. MARCADORES MOLECULARES DERIVADOS DE SEQÜÊNCIAS EXPRESSAS DO GENOMA DO CAFÉ POTENCIALMENTE ENVOLVIDAS NA RESISTÊNCIA À FERRUGEM. Universidade Federal de Vicosa. 2007.

27. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. J Comput Biol. 2012;19: 455–477. doi:10.1089/cmb.2012.0021

28. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013;14: R36.

doi:10.1186/gb-2013-14-4-r36

29. Stanke M, Steinkamp R, Waack S, Morgenstern B. AUGUSTUS: a web server for gene finding in eukaryotes. Nucleic Acids Res. 2004;32: W309–W312. doi:10.1093/nar/gkh379

30. Lin C, Mueller LA, Carthy JM, Crouzillat D, Pétiard V, Tanksley SD. Coffee and tomato share common gene repertoires as revealed by deep sequencing of seed and cherry transcripts. Theor Appl Genet. 2005;112: 114–130. doi:10.1007/s00122-005-0112-2

31. Yachdav G, Kloppmann E, Kajan L, Hecht M, Goldberg T, Hamp T, et al. PredictProtein--an open resource for online prediction of protein structural and functional features. Nucleic Acids Res. 2014;42: W337–W343. doi:10.1093/nar/gku366

32. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. Nat Methods. 2014;12: 7–8. doi:10.1038/nmeth.3213

33. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nat Biotechnol. 2011;29: 24–26. doi:10.1038/nbt.1754

34. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016;33: 1870–1874. doi:10.1093/molbev/msw054

35. Rozas J. DNA sequence polymorphism analysis using DnaSP. Methods in molecular biology (Clifton, N.J.). 2009. pp. 337–350. doi:10.1007/978-1-59745-251-9_17

36. Rzhetsky a, Nei M. A Simple Method for Estimating and Testing Minimum-Evolution Trees. Mol Biol Evol. 1992;9: 945–967.

37. Meyers BC, Kaushik S, Nandety RS. Evolving disease resistance genes. Curr Opin Plant Biol. 2005;8: 129–134. doi:10.1016/j.pbi.2005.01.002

38. Hammond-Kosack KE, Kanyuka K. Resistance Genes (R Genes) in Plants. eLS. Chichester, UK: John Wiley & Sons, Ltd; 2007. doi:10.1002/9780470015902.a0020119

39. Goyal N, Bhatia G, Sharma S, Garewal N, Upadhyay A, Upadhyay SK, et al. Genome-wide

characterization revealed role of NBS-LRR genes during powdery mildew infection in Vitis vinifera. Genomics. 2019; doi:10.1016/j.ygeno.2019.02.011

40. Mur LAJ, Kenton P, Lloyd AJ, Ougham H, Prats E. The hypersensitive response; the centenary is upon us but how much do we know? J Exp Bot. 2008;59: 501–520. doi:10.1093/jxb/erm239

41. Sanabria N, Goring D, Nürnberger T, Dubery I. Self/nonself perception and recognition mechanisms in plants: a comparison of self-incompatibility and innate immunity. New Phytol. 2008;178: 503–514. doi:10.1111/j.1469-8137.2008.02403.x

42. Qi D, Innes RW. Recent Advances in Plant NLR Structure, Function, Localization, and Signaling. Front Immunol. 2013;4. doi:10.3389/fimmu.2013.00348

43. Rairdan GJ, Collier SM, Sacco MA, Baldwin TT, Boettrich T, Moffett P. The Coiled-Coil and Nucleotide Binding Domains of the Potato Rx Disease Resistance Protein Function in Pathogen Recognition and Signaling. PLANT CELL ONLINE. 2008;20: 739–751. doi:10.1105/tpc.107.056036

44. Kim SH, Kwon SI, Saha D, Anyanwu NC, Gassmann W. Resistance to the Pseudomonas syringae Effector HopA1 Is Governed by the TIR-NBS-LRR Protein RPS6 and Is Enhanced by Mutations in SRFR1. PLANT Physiol. 2009;150: 1723–1732. doi:10.1104/pp.109.139238

45. Ratnaparkhe MB, Wang X, Li J, Compton RO, Rainville LK, Lemke C, et al. Comparative analysis of peanut NBS-LRR gene clusters suggests evolutionary innovation among duplicated domains and erosion of gene microsynteny. New Phytol. 2011;192: 164–178. doi:10.1111/j.1469-8137.2011.03800.x

46. Djebbi S, Bouktila D, Makni H, Makni M, Mezghani-Khemakhem M. Identification and characterization of novel NBS-LRR resistance gene analogues from the pea. Genet Mol Res. 2015;14: 6419–6428. doi:10.4238/2015.June.11.18

47. Bettencourt J. Considerações gerais sobre o “Híbrido de Timor” [Internet]. 1st ed. Journal of Chemical Information and Modeling. Sao Paulo: Instituto Agronômico; 1973. doi:10.1017/CBO9781107415324.004

48. Noronha-Wagner and Bettencourt A. Genetic study of the resistance of Coffea sp to leaf rust 1. Identification and behavior of four factors conditioning disease reaction in Coffea arabica to twelve physiologic races of Hemileia vastatrix. Can J Bot. 1967;45: 2021–2031.

49. Bettencourt, Rodrigues. Principles and practice of coffee breeding for resistance to rust and other diseases. Elsevier Appl Sci. 1988;3: 199–234.

50. Bettencourt AJ, Coronha-Wagner. Genetic factors conditioning resistance of Coffea arabica L. to Hemileia vastatrix Berk et Br. Agron Lusit. 1971;31: 285–292.

51. Kang Y, Kim K, Shim S, Yoon M, Sun S, Kim M, et al. Genome-wide mapping of NBS-LRR genes and their association with disease resistance in soybean. BMC Plant Biol. 2012;12: 139. doi:10.1186/1471-2229-12-139

52. González VM, Aventín N, Centeno E, Puigdomènech P. Interspecific and intraspecific gene variability in a 1-Mb region containing the highest density of NBS-LRR genes found in the melon genome. BMC Genomics. 2014;15: 1131. doi:10.1186/1471-2164-15-1131

53. Sanseverino W, Hénaff E, Vives C, Pinosio S, Burgos-Paz W, Morgante M, et al. Transposon Insertions, Structural Variations, and SNPs Contribute to the Evolution of the Melon Genome. Mol Biol Evol. 2015;32: 2760–2774. doi:10.1093/molbev/msv152

54. Panchy N, Lehti-Shiu MD, Shiu S-H. Evolution of gene duplication in plants. Plant Physiol. 2016; pp.00523.2016. doi:10.1104/pp.16.00523

55. van der Biezen EA, Jones JD. The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. Curr Biol. 1998;8: R226-7. doi:10.1016/S0960-9822(98)70145-9 showArticle Info

56. Wang G-F, Ji J, EI-Kasmi F, Dangl JL, Johal G, Balint-Kurti PJ. Molecular and Functional Analyses of a Maize Autoactive NB-LRR Protein Identify Precise Structural Requirements for Activity. Mackey D, editor. PLOS Pathog. 2015;11: e1004674. doi:10.1371/journal.ppat.1004674

57. Hao W, Collier SM, Moffett P, Chai J. Structural Basis for the Interaction between the Potato Virus X

Resistance Protein (Rx) and Its Cofactor Ran GTPase-activating Protein 2 (RanGAP2). J Biol Chem. 2013;288: 35868–35876. doi:10.1074/jbc.M113.517417

58. Lahaye T. The Arabidopsis RRS1-R disease resistance gene--uncovering the plant's nucleus as the new battlefield of plant defense? Trends Plant Sci. 2002;7: 425–7. Available: http://www.ncbi.nlm.nih.gov/pubmed/12399170

59. Kierszniowska S, Seiwert B, Schulze WX. Definition of Arabidopsis Sterol-rich Membrane Microdomains by Differential Treatment with Methyl- -cyclodextrin and Quantitative Proteomics. Mol Cell Proteomics. 2009;8: 612–623. doi:10.1074/mcp.M800346-MCP200

60. Gou X, He K, Yang H, Yuan T, Lin H, Clouse SD, et al. Genome-wide cloning and sequence analysis of leucine-rich repeat receptor-like protein kinase genes in Arabidopsis thaliana. BMC Genomics. 2010;11: 19. doi:10.1186/1471-2164-11-19

61. Xu Y, Liu F, Zhu S, Li X. The Maize NBS-LRR Gene ZmNBS25 Enhances Disease Resistance in Rice and Arabidopsis. Front Plant Sci. 2018;9. doi:10.3389/fpls.2018.01033

62. Michelmore RW, Meyers BC. Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. Genome Res. 1998;8: 1113–30. Available: http://www.ncbi.nlm.nih.gov/pubmed/9847076

63. Padmanabhan M, Cournoyer P, Dinesh-Kumar SP. The leucine-rich repeat domain in plant innate immunity: a wealth of possibilities. Cell Microbiol. 2009;11: 191–198. doi:10.1111/j.1462-5822.2008.01260.x

64. Zhao Y, Huang J, Wang Z, Jing S, Wang Y, Ouyang Y, et al. Allelic diversity in an NLR gene BPH9 enables rice to combat planthopper variation. Proc Natl Acad Sci. 2016;113: 12850–12855. doi:10.1073/pnas.1614862113

65. Velasco R, Zharkikh A, Troggio M, Cartwright DA, Cestaro A, Pruss D, et al. A High Quality Draft Consensus Sequence of the Genome of a Heterozygous Grapevine Variety. Dilkes B, editor. PLoS One. 2007;2: e1326. doi:10.1371/journal.pone.0001326

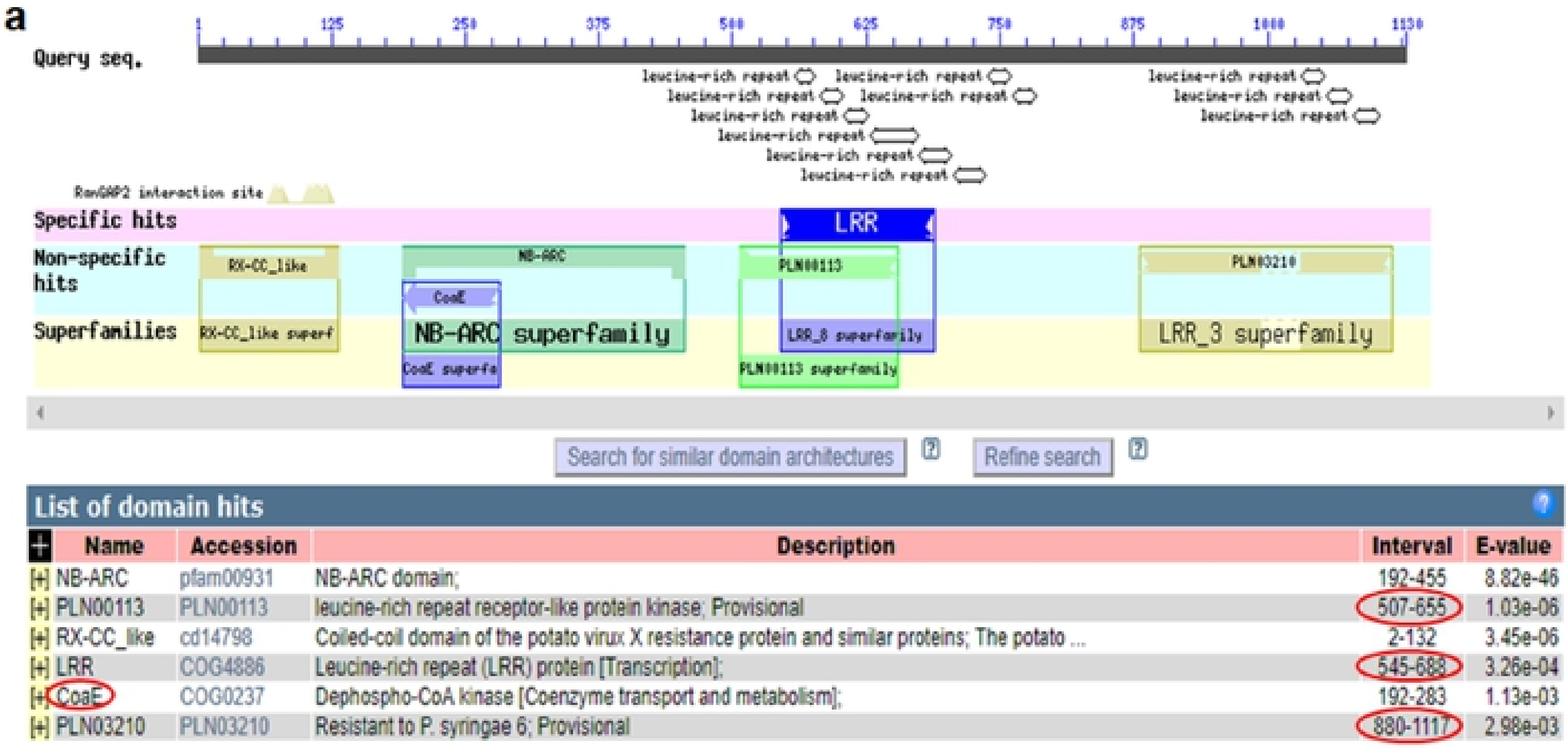
66. Sara M, Walter S, Paola C, Luigi M, Luigi F, Raffalella EM. Solanaceae Evolutionary Dynamics of the &amp;lt;i&amp;gt;I&amp;lt;/i&amp;gt;2-NBS Domain. Am J Plant Sci. 2012;03: 283–294. doi:10.4236/ajps.2012.32034

67. Felsenstein J. Confidence-Limits on Phylogenies - an Approach Using the Bootstrap. Evolution (N Y). 1985; doi:Doi 10.2307/2408678

68. Zuckerkandl E, Pauling L. Evolutionary divergence and convergence in proteins. Evol genes proteins. 1965; 97–166. doi:10.1209/epl/i1998-00224-x

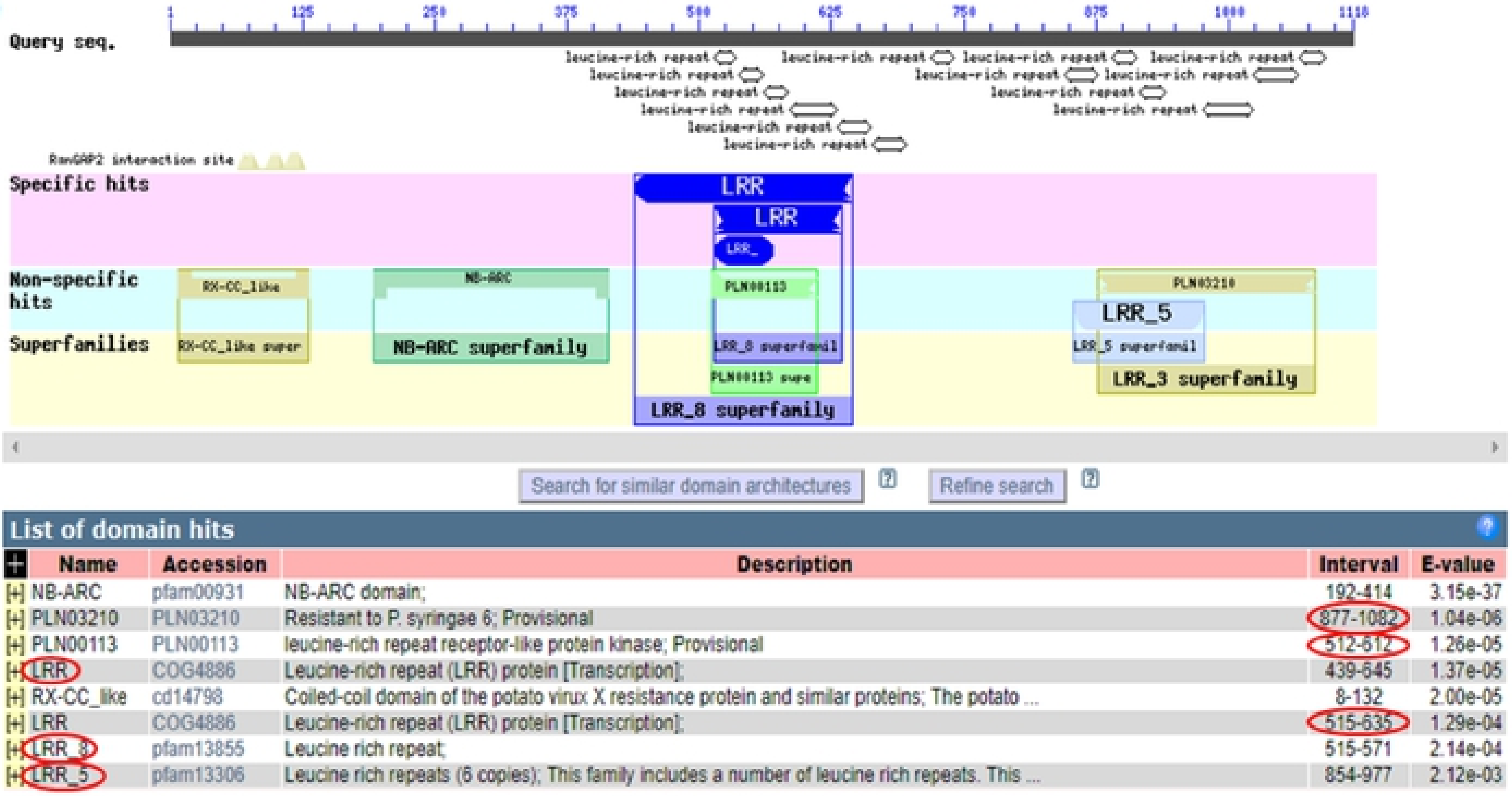
69. Nei M, Kumar S. Molecular evolution and phylogenetics. 1st ed. Oxford: Oxford University Press; 2000.

70. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. Oxford University Press; 1987;4: 406–25. Available: http://www.ncbi.nlm.nih.gov/pubmed/3447015



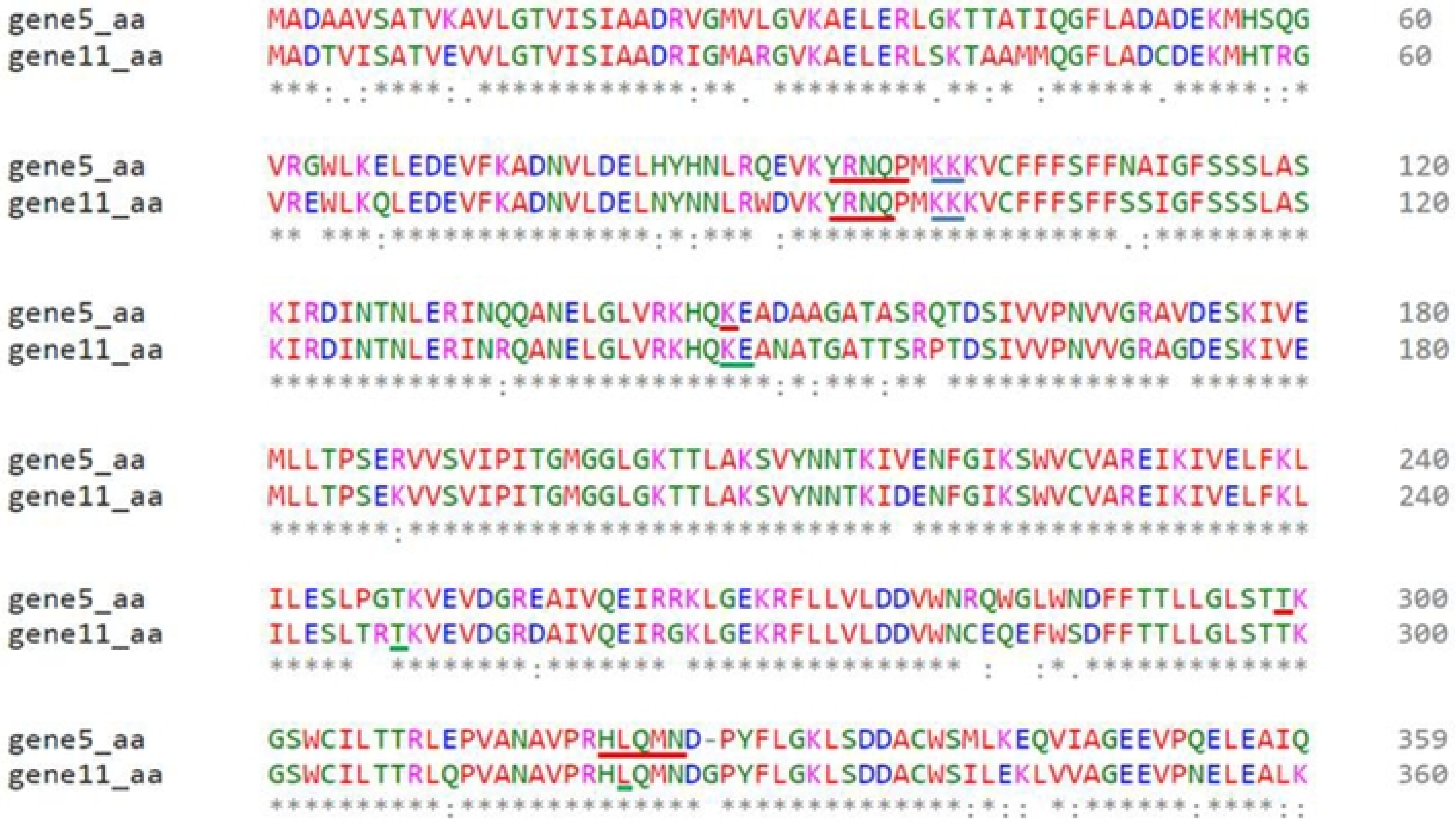
Name	Accession	Description	Interval	E-value
[+] NB-ARC	pfam00931	NB-ARC domain;	192-455	8.82e-46
[+] PLN00113	PLN00113	leucine-rich repeat receptor-like protein kinase; Provisional	507-655	1.03e-06
[+] RX-CC_like	cd14798	Coiled-coil domain of the potato virux X resistance protein and similar proteins; The potato ...	2-132	3.45e-06
[+] LRR	COG4886	Leucine-rich repeat (LRR) protein [Transcription];	545-688	3.26e-04
[+] CoaE	COG0237	Dephospho-CoA kinase [Coenzyme transport and metabolism];	192-283	1.13e-03
[+] PLN03210	PLN03210	Resistant to P. syringae 6; Provisional	880-1117	2.98e-03

b

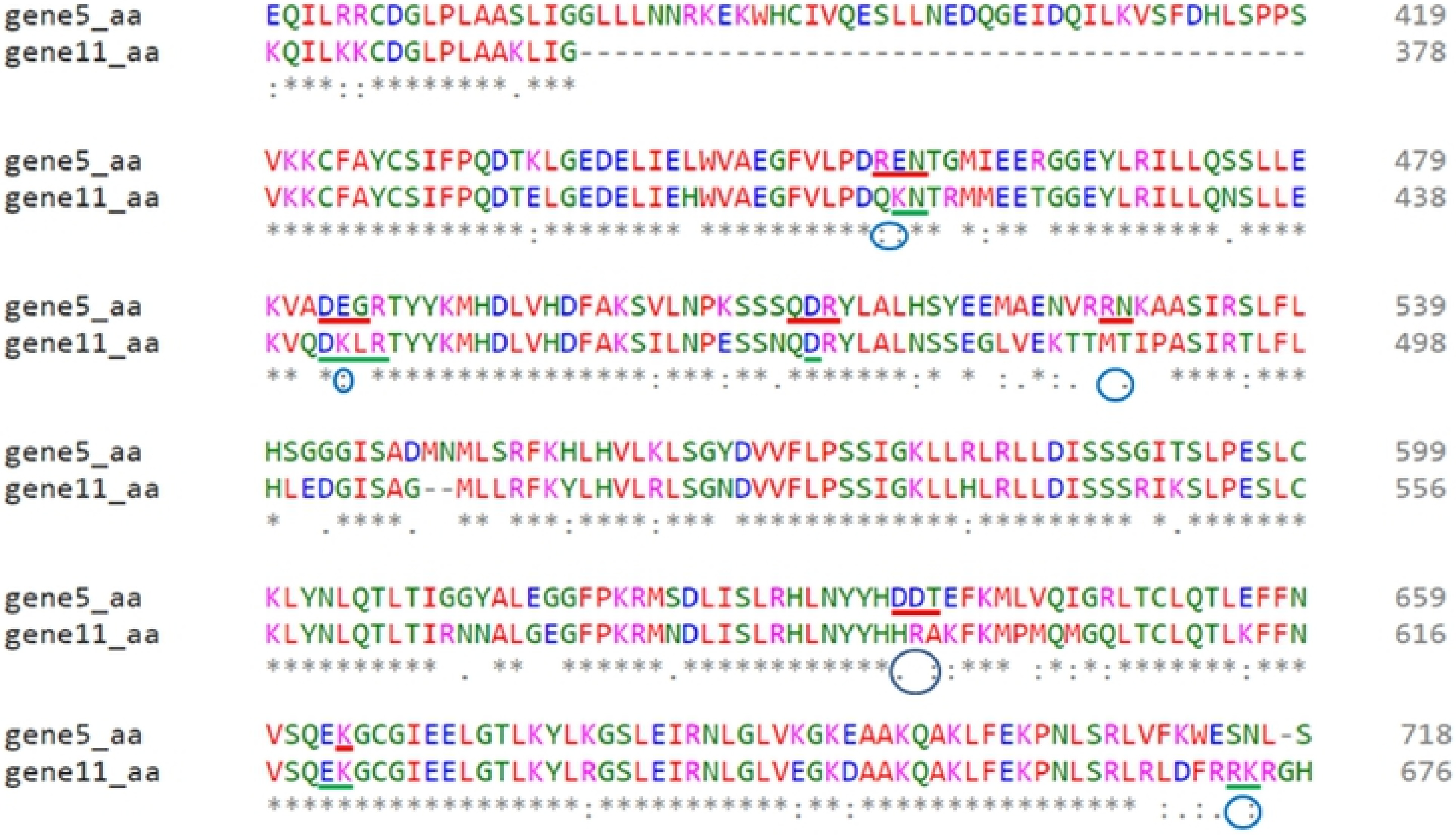


List of domain hits

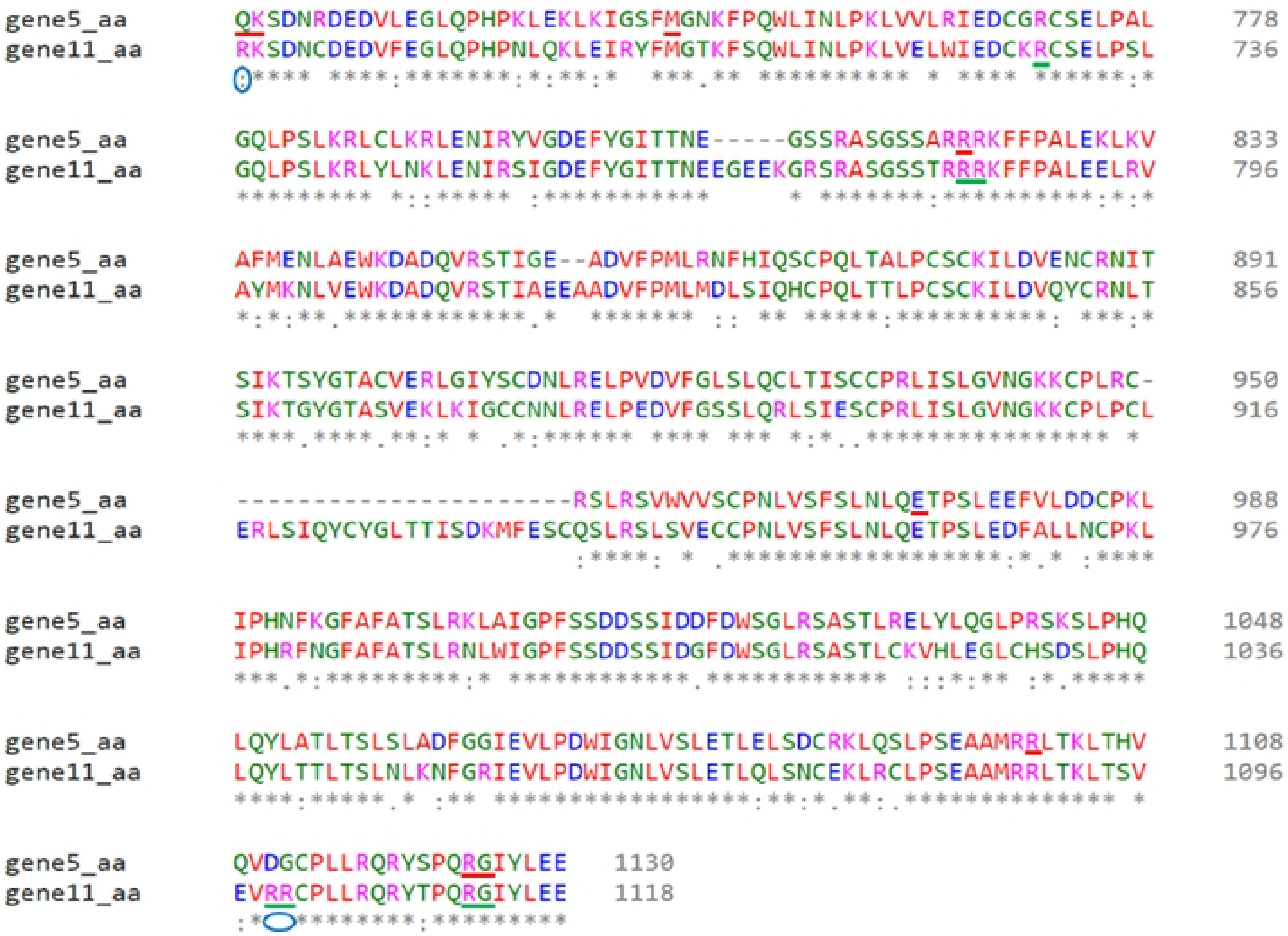
+	Name	Accession	Description	Interval	E-value
[+]	NB-ARC	pfam00931	NB-ARC domain;	192-414	3.15e-37
[+]	PLN03210	PLN03210	Resistant to P. syringae 6; Provisional	877-1082	1.04e-06
[+]	PLN00113	PLN00113	leucine-rich repeat receptor-like protein kinase; Provisional	512-612	1.26e-05
[+]	LRR	COG4886	Leucine-rich repeat (LRR) protein [Transcription];	439-645	1.37e-05
[+]	RX-CC_like	cd14798	Coiled-coil domain of the potato virus X resistance protein and similar proteins; The potato ...	8-132	2.00e-05
[+]	LRR	COG4886	Leucine-rich repeat (LRR) protein [Transcription];	515-635	1.29e-04
[+]	LRR_8	pfam13855	Leucine rich repeat;	515-571	2.14e-04
[+]	LRR_5	pfam13306	Leucine rich repeats (6 copies); This family includes a number of leucine rich repeats. This ...	854-977	2.12e-03



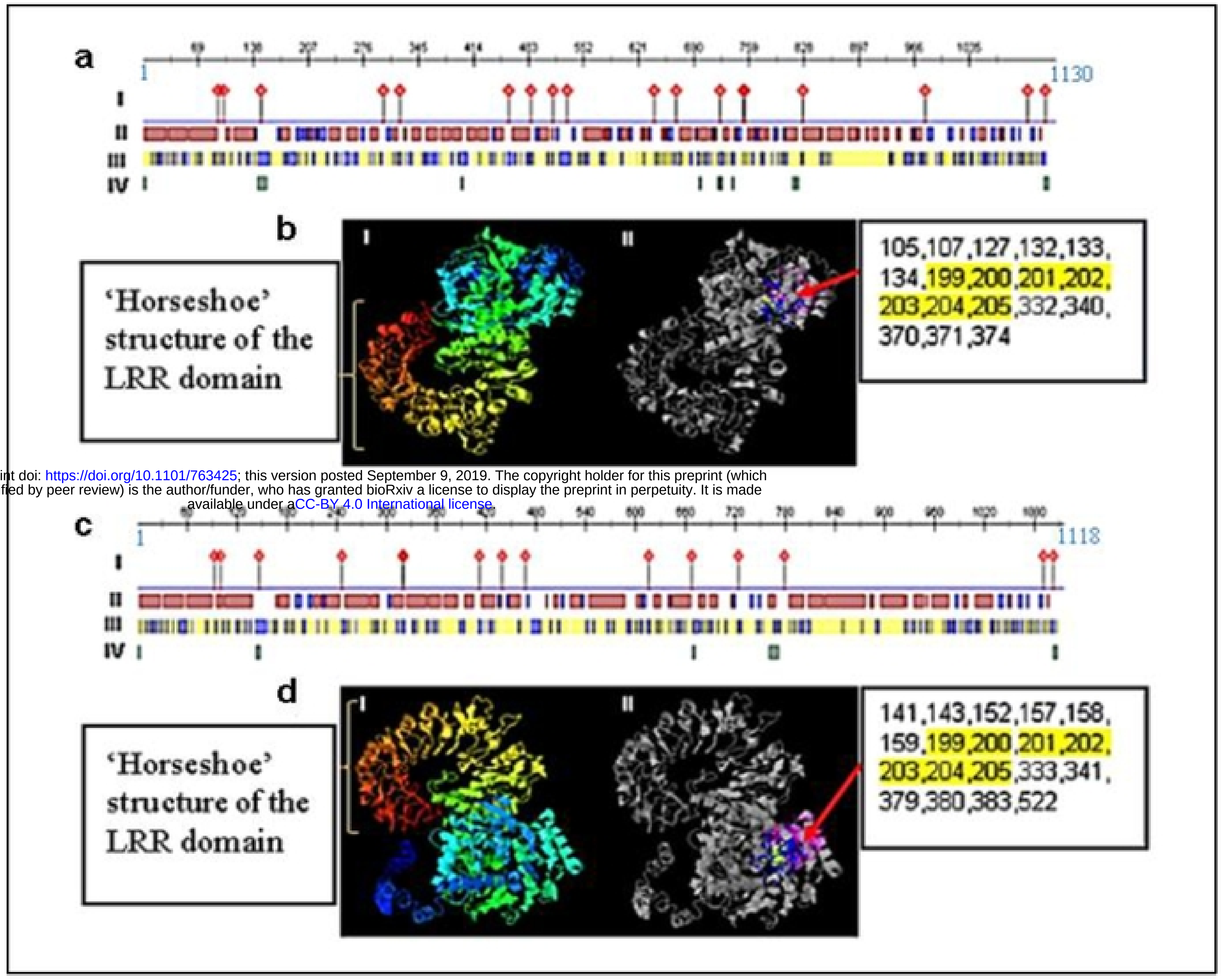
```
gene5_aa        MADAAVSATVKAVLGTVISIAADRVGMVLGVKAELERLGKTTATIQGFLADADEKMHSQG    60
gene11_aa       MADTVISATVEVVLGTVISIAADRIGMARGVKAELERLSKTAAMMQGFLADCDEKMHTRG    60
                ***:.:****:.************:**. *********.**:* :******.*****::*

gene5_aa        VRGWLKELEDEVFKADNVLDELHYHNLRQEVKYRNQPMKKKVCFFFSFFNAIGFSSSLAS    120
gene11_aa       VREWLKQLEDEVFKADNVLDELNYNNLRWDVKYRNQPMKKKVCFFFSFFSSIGFSSSLAS    120
                ** ***:***************:*:*** :*******************.:*********

gene5_aa        KIRDINTNLERINQQANELGLVRKHQKEADAAGATASRQTDSIVVPNVVGRAVDESKIVE    180
gene11_aa       KIRDINTNLERINRQANELGLVRKHQKEANATGATTSRPTDSIVVPNVVGRAGDESKIVE    180
                *************:***************:*:***:** ************* *******

gene5_aa        MLLTPSERVVSVIPITGMGGLGKTTLAKSVYNNTKIVENFGIKSWVCVAREIKIVELFKL    240
gene11_aa       MLLTPSEKVVSVIPITGMGGLGKTTLAKSVYNNTKIDENFGIKSWVCVAREIKIVELFKL    240
                *******:**************************** ***********************

gene5_aa        ILESLPGTKVEVDGREAIVQEIRRKLGEKRFLLVLDDVWNRQWGLWNDFFTTLLGLSTIK    300
gene11_aa       ILESLTRTKVEVDGRDAIVQEIRGKLGEKRFLLVLDDVWNCEQEFWSDFFTTLLGLSTTK    300
                *****  ********:******* **************** :   :*.*************

gene5_aa        GSWCILTTRLEPVANAVPRHLQMND-PYFLGKLSDDACWSMLKEQVIAGEEVPQELEAIQ    359
gene11_aa       GSWCILTTRLQPVANAVPRHLQMNDGPYFLGKLSDDACWSILEKLVVAGEEVPNELEALK    360
                **********:************** **************:*:: *:******:****::
```



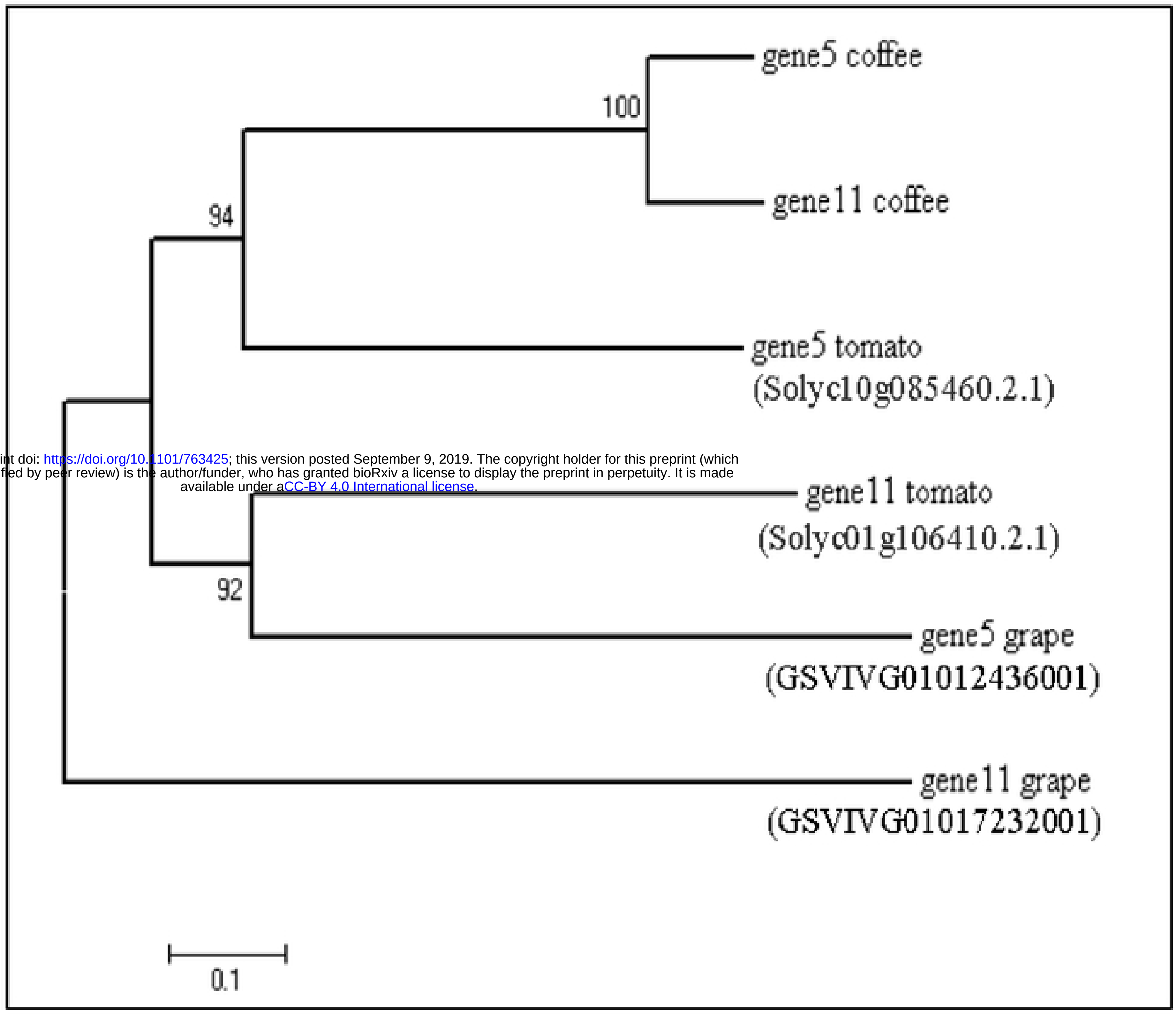
```
gene5_aa        EQILRRCDGLPLAASLIGGLLLNNRKEKWHCIVQESLLNEDQGEIDQILKVSFDHLSPPS     419
gene11_aa       KQILKKCDGLPLAAKLIG------------------------------------------     378
                :***::********.***

gene5_aa        VKKCFAYCSIFPQDTKLGEDELIELWVAEGFVLPDRENTGMIEERGGEYLRILLQSSLLE     479
gene11_aa       VKKCFAYCSIFPQDTELGEDELIEHWVAEGFVLPDQKNTRMMEETGGEYLRILLQNSLLE     438
                ***************:******** **********: ** *:** **********.****

gene5_aa        KVADEGRTYYKMHDLVHDFAKSVLNPKSSSQDRYLALHSYEEMAENVRRNKAASIRSLFL     539
gene11_aa       KVQDKLRTYYKMHDLVHDFAKSILNPESSNQDRYLALNSSEGLVEKTTMTIPASIRTLFL     498
                ** *: ****************:***:**.*******:* * :.*:.  .  ****:***

gene5_aa        HSGGGISADMNMLSRFKHLHVLKLSGYDVVFLPSSIGKLLRLRLLDISSSGITSLPESLC     599
gene11_aa       HLEDGISAG--MLLRFKYLHVLRLSGNDVVFLPSSIGKLLHLRLLDISSSRIKSLPESLC     556
                *  .****.  ** ***:****:*** *************:********* *.*******

gene5_aa        KLYNLQTLTIGGYALEGGFPKRMSDLISLRHLNYYHDDTEFKMLVQIGRLTCLQTLEFFN     659
gene11_aa       KLYNLQTLTIRNNALGEGFPKRMNDLISLRHLNYYHHRAKFKMPMQMGQLTCLQTLKFFN     616
                ********** . **  ******.************. ::*** :*:*:*******:***

gene5_aa        VSQEKGCGIEELGTLKYLKGSLEIRNLGLVKGKEAAKQAKLFEKPNLSRLVFKWESNL-S     718
gene11_aa       VSQEKGCGIEELGTLKYLRGSLEIRNLGLVEGKDAAKQAKLFEKPNLSRLRLDFRRKRGH     676
                ******************:***********:**:**************** :.:. :
```



```
gene5_aa     QKSDNRDEDVLEGLQPHPKLEKLKIGSFMGNKFPQWLINLPKLVVLRIEDCGRCSELPAL   778
gene11_aa    RKSDNCDEDVFEGLQPHPNLQKLEIRYFMGTKFSQWLINLPKLVELWIEDCKRCSELPSL   736
             :**** ****:*******:*:**:*  ***.** ********** * **** ******:*

gene5_aa     GQLPSLKRLCLKRLENIRYVGDEFYGITTNE-----GSSRASGSSARRRKFFPALEKLKV   833
gene11_aa    GQLPSLKRLYLNKLENIRSIGDEFYGITTNEEGEEKGRSRASGSSTRRRKFFPALEELRV   796
             ********* *::***** :***********     * *******:**********:*:*

gene5_aa     AFMENLAEWKDADQVRSTIGE--ADVFPMLRNFHIQSCPQLTALPCSCKILDVENCRNIT   891
gene11_aa    AYMKNLVEWKDADQVRSTIAEEAADVFPMLMDLSIQHCPQLTTLPCSCKILDVQYCRNLT   856
             *:*:**.************.*  ******* :: ** *****:**********: ***:*

gene5_aa     SIKTSYGTACVERLGIYSCDNLRELPVDVFGLSLQCLTISCCPRLISLGVNGKKCPLRC-   950
gene11_aa    SIKTGYGTASVEKLKIGCCNNLRELPEDVFGSSLQRLSIESCPRLISLGVNGKKCPLPCL   916
             ****.****.**:* * .*:****** **** *** *:*..**************** *

gene5_aa     ----------------------RSLRSVWVVSCPNLVSFSLNLQETPSLEEFVLDDCPKL   988
gene11_aa    ERLSIQYCYGLTTISDKMFESCQSLRSLSVECCPNLVSFSLNLQETPSLEDFALLNCPKL   976
                                   :****: * .******************:*.* :****

gene5_aa     IPHNFKGFAFATSLRKLAIGPFSSDDSSIDDFDWSGLRSASTLRELYLQGLPRSKSLPHQ  1048
gene11_aa    IPHRFNGFAFATSLRNLWIGPFSSDDSSIDGFDWSGLRSASTLCKVHLEGLCHSDSLPHQ  1036
             ***.*:*********:* ************.************ :::*:** :*.*****

gene5_aa     LQYLATLTSLSLADFGGIEVLPDWIGNLVSLETLELSDCRKLQSLPSEAAMRRLTKLTHV  1108
gene11_aa    LQYLTTLTSLNLKNFGRIEVLPDWIGNLVSLETLQLSNCEKLRCLPSEAAMRRLTKLTSV  1096
             ****:*****.* :** *****************:**:*.**:.************** *

gene5_aa     QVDGCPLLRQRYSPQRGIYLEE   1130
gene11_aa    EVRRCPLLRQRYTPQRGIYLEE   1118
             :*  ********:*********
```


a
69
138
207
276
345
414
483
552
621
690
759
828
897
966
1035
1
1130
I
II
III
IV
b
I
II
'Horseshoe' structure of the LRR domain
105,107,127,132,133, 134,199,200,201,202, 203,204,205,332,340, 370,371,374
c
60
120
180
240
300
360
420
480
540
600
660
720
780
840
900
960
1020
1080
1
1118
I
II
III
IV
d
I
II
'Horseshoe' structure of the LRR domain
141,143,152,157,158, 159,199,200,201,202, 203,204,205,333,341, 379,380,383,522

Figure


gene5 coffee
100
gene11 coffee
94
gene5 tomato
(Solyc10g085460.2.1)
gene11 tomato
(Solyc01g106410.2.1)
92
gene5 grape
(GSVIVG01012436001)
gene11 grape
(GSVIVG01017232001)
0.1

Figure